(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 434 415 A2

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 90313970.7

(51) Int. Cl.^{5}: G06F 15/70

(22) Date of filing: 20.12.90

(30) Priority: 21.12.89 US 454339

(43) Date of publication of application: 26.06.91 Bulletin 91/26

(84) Designated Contracting States: DE FR GB

(71) Applicant: XEROX CORPORATION
Xerox Square
Rochester New York 14644 (US)

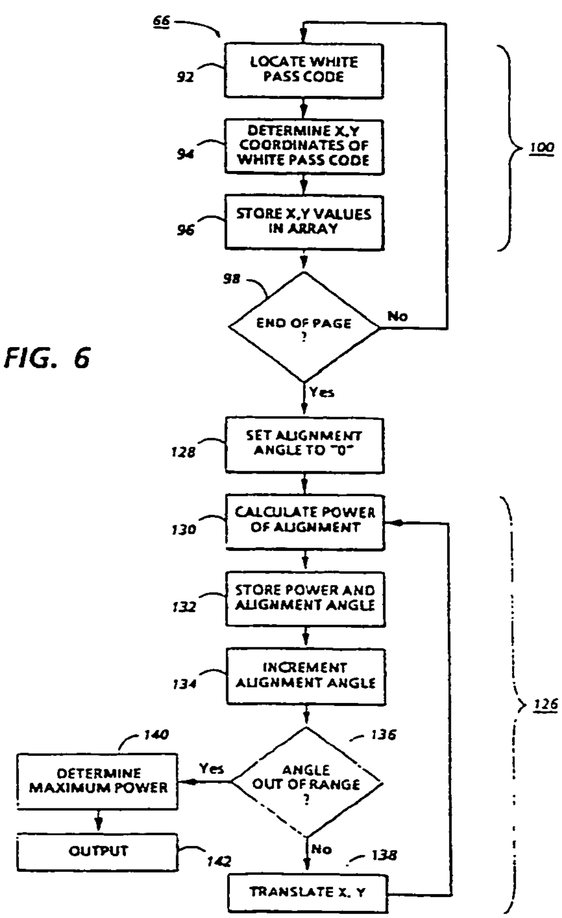
(72) Inventor: Spitz, A. Lawrence
2311 Bryant Street
Palo Alto, California, 94301 (US)

(74) Representative: Weatherald, Keith Baynes et al
Rank Xerox Patent Department Albion House,
55 New Oxford Street
London WC1A 1BS (GB)

(54) Method of measuring skew angles.

(57) Skew angle of an image is determined based on determination of location of fiducial points on the image. Fiducial points may be located through a comparison of the scanning of a first line with scanning of a subsequent line. These fiducial points may be defined in terms of pixel color transitions located on a first scan line without a corresponding transition on the succeeding scan line. Skew angle may be determined from image data in uncompressed form or in compressed form. Where skew angle is determined from image data in compressed form, the two-dimensional CCITT facsimile recommendations may be used. In such cases, the locations of the fiducial points may be taken as the locations of the pass codes of the compressed image data. Specifically, pass codes indicating a pass of white pixels are used.

FIG. 6



EP 0 434 415 A2

of the present invention) to provide a means for recognizing and compensating for their effect on determination of the true baseline. Note that there is exactly one fiducial point per connected component in the image.

In essence, Baird calculates skew by determining the number of fiducial points per "line" for a variety of rotational alignments. 'Line' as used herein means, for example, one of a plurality of imaginary parallel scan lines traversing the document and oriented perpendicular to a selected feature such as a margin or page edge. The rotational alignments are calculated by trigonometric translation of the fiducial points.

Referring to Figs. 1b and 1c of the drawings, counting of the number of fiducial points per line is accomplished by projecting the locations of the fiducial points 10 onto an accumulator line 18 which is perpendicular to the projection direction, as indicated by arrow p. Accumulator line 18 is partitioned into "bins" 20 of a uniform predetermined height, H, for example equal to 1/3 of the height of a six-point character. Height H may be varied as appropriate, and may be as small as two pixels. However, as H decreases, computation time increases. Importantly, as H approaches the character height, skew angle determination performance disintegrates. Returning to the bins, there is exactly one bin per line. The number of fiducial points for a selected line is then equal to the number of fiducial points projected into the bin corresponding to that line.

Since this method results in a relatively small number of fiducial points (depending on the nature of the image), the alignment is made efficient by calculating the alignment on the basis of the sum of a positive power greater than 1, e.g., two (sum of squares) of the counts of the fiducial points which appear in each of the rotationally-aligned bins. The variance of the distribution is maximized by maximizing the sum of squares of the counts, resulting in an index of the "power" of the alignment, from which skew angle is determined. Fig. 1b shows the positions of the fiducial points 10 and the relative size of each bin 20 on a skewed text sample in which the bins are unaligned with the skewed text. Fig. 1c shows the distribution of the fiducial points of the same skewed image of Fig. 1b into aligned bins.

Calculation of the power of each of a variety of alignments requires that the positions of each fiducial point be known. Specifically, the coordinates of a fiducial point are used to translate the fiducial point mathematically, by an angle and a displacement from an origin, to a new set of coordinates. This process is done for the complete collection of fiducial points, and the powers of the alignments before and after translation are compared. From each comparison, the angle corresponding to the alignment with the greatest power is retained. After all angular alignments within a selected range have been compared in this manner, the skew angle may be assumed to be the angle corresponding to the alignment with the greatest power. In the event of weak alignment or multiple alignments, however, this assumption may need to be verified otherwise.

The present invention provides a method of skew angle determination, overcoming a number of known problems and shortcomings.

One aspect of the present invention is the discovery of the problem that, in Baird's method of skew angle determination, as skew angle is increased, the distance between the bottom of the mark and the bottom of the bounding box may increase. This difference increases as a function of the sine of the angle of skew. This leads to the problem of errors in the count of fiducial points, which affects skew angle determination as a square of that error.

This aspect is further based on the discovery that this problem can be solved by determining skew angle based on fiducial points located on the mark itself. In this manner, the extent of the skew angle does not contribute to error in the determination of that angle.

Related to this aspect of the present invention is the aspect that it has been discovered that line ends form a useful set of features on which to base skew determination. A particularly useful set of features are the line ends located at or near a baseline for horizontally-oriented images or a vertical line for vertical or columnar-oriented images. Detecting line ends in generic images has proven to be a difficult task. However, a method has been invent, and which forms another aspect of the present invention, to detect a group of features that includes the group of line ends. The set of features is limited in size such that the effectiveness of using line ends is not diminished. By comparing one scan line to a subsequent scan line, pixel color transitions may be located. Proper selection of the pixel color transitions can yield a set of features, including lines ends, which can be used very effectively to determine skew angle.

Another aspect of the present invention is the recognition that another problem of the prior art methods of skew angle determination is that they involve computationally demanding methods of feature (fiducial point) location. According to Baird's method, not only must the mark be identified, but a bounding box must be constructed around the mark, and the bottom center of the bounding box must be located.

This problem can be solved by determining fiducial points from a comparison of the scanning of a first line with scanning of a subsequent line. Location of fiducial points will correspond to the location of selected topographic features located on a first scan line without a corresponding topographic feature on the succeeding scan line or, alternatively, without a corresponding topographic feature located on the preceding scan line. In this regard, for the purposes of this disclosure, "topographic" feature is taken to mean a feature of a mark, image

Fig. 6 shows a flow diagram of one aspect of the present invention ;
Fig. 7 shows a flow diagram of another aspect of the present invention ;
Fig. 8 shows a portion of skewed text used as an example of one aspect of the present invention ;
Fig. 9 shows the locations of the white pass codes for the sample text of Fig. 8 used to determime alignment according to the present invention ;
Fig. 10 shows a plot of the power of the alignments for an angle range for the sample skewed text of Fig. 8, and
Fig. 11 shows a plot of power of alignments for text having multiple skew angles.

The method of skew determination according to the present invention will now be described. It is particularly applicable to determination of skew in scanned images. In the present disclosure it is assumed that the image data are broken up into distinct units, preferably lines of the image. This may be accomplished, for example, by scanning the image and inserting an indication of the beginning and ending of each scan line or of the run length, as is well known. For convenience these units are hereafter referred to as 'scan lines'.

With increasing prevalence, scanned images are being handled in compressed format. For example, facsimile machines operate on scanned images virtually entirely in compressed format. For this reason, the present invention is described assuming that the image data are in compressed form (also referred to herein as the compressed data domain). It will be evident to one skilled in the art, however, that the present invention is equally applicable to image data in uncompressed format.

Fig. 2 illustrates in block format an environment in which the present invention may operate. Specifically, Fig. 2 illustrates a portion of a computer system 50 tailored to operate according to the present invention. Computer system 50 includes or is connected to receive output signals from a scanner 52, which is capable of scanning an image and producing digital data which represent that image. These digital data are communicated to a processor 54. This processor controls input and output operations and calls to program memory 56 and data memory 58 *via* bus 60.

Program memory 56 may include, *inter alia*, a routine 62 for controlling the scanning of an image by scanner 62, a routine 64 for converting the digital data representing the image into a compressed data format, and a routine 66 for determining skew angle from the compressed data. Program memory 56 will thus have associated with it data memory 58 in which may be stored, *inter alia*, at location 68 the digital data structure produced by scanner 52 under control of the scanning control routine 62, at location 70 the data structure of the compressed representation of the scanned image produced by compression routine 64, and at location 72 the data structure containing selected point data, for example fiducial point location, produced by skew angle determination routine 66. To facilitate the communication between program memory 56 and data memory 58 necessary for operation, each is connected to bus 60 such that input and output operations may be performed. One additional point with regard to memories 56 and 58 is that they have been described as separate for the purposes of clarity but they, in fact, be parts of a single memory block of the computer system.

Under processor control, skew detection routine 66 will access various parts of data memory 58 to acquire data needed to calculate skew angle. Once calculated, skew angle may then be output at 74, which may comprise a means for displaying the results, such as a CRT display, printer or the like, or may comprise a means for utilizing the results to perform further operations, such as modification of the image data to compensate for skewness, etc.

It has been assumed that the image data have been compressed according to the Group 4 standard, although the present invention with proper modification will render similar results using other compression schemes, e.g., CCITT two-dimensional Group 3 format, etc. The coding scheme of Group 4 relies on the existence and relative spacing between pixel color transitions found on pairs of succeeding scan lines. In Group 4 coding each line in turn becomes a "coding line" and is coded with respect to its predecessor, the "reference line". The first line is coded with respect to an artificially defined all-white reference line. See Hunter, *et al.*, "International Digital Facsimile Coding Standards," Proceedings of the IEEE, Vol. 68, No. 7, July 1980, pp. 854-867 and Int'l. Telecommunications Union, CCITT (Int'l. Telegraph and Telephone Consultative Committee) Red Book, Geneva 1985 (ISBN 92-61-02291-X) for a more detailed discussion of the Group 4 compression standard.

Encoding in the Group 4 format has three modes – vertical, horizontal and pass. These modes are described following with regard to Figs. 3a, 3b, and 3c. Adjacent scan lines are compared to determine whether, given a first pixel color transition on a reference line such as black to white, there exists a corresponding pixel color transition (i.e., also black to white) on the coding line. The existence and relative spacing of the transition on the coding line from the transition on the reference line is used to determine the mode.

With reference to Fig. 3a, a vertical mode is used when the black-to-white or white-to-black transition positions on adjacent scan lines are horizontally close (≤ 3 pixels) and thus can be encoded in a small number of bits. Horizontal mode is used when the transition positions are further apart than three pixels, as shown in Fig.

at location 72 of data memory 58, are located. This is indicated by box 92 of Fig. 6. Once a white pass code is located, its location in an appropriate coordinate system must be determined, as done by box 94. Rectangular coördinates having an abscissa, x, and an ordinate, y, coordinate pairs are generally convenient for this purpose, although other coordinate systems, such as polar coordinates, may be used when appropriate. An array of the coordinate pairs may then be constructed at 96 for use in computing alignments, as discussed in detail below. After storing the coordinate pairs of the locations of the white pass codes, a test is performed at 98 to determine whether the end of the scanned page has been reached. If so, skew angle determination proceeds as discussed in detail below. If not, a search is made for the next, if any, white pass code on the page.

The previous four steps 92 through 98 are collectively referred to as coordinate determination routine 100, which may be a subroutine of skew determination routine 66, and which is further described now with reference to Fig. 7. Block 101 illustrates an input of data in the Group 4 compressed format. Using x-y coordinate pairs, x and y must first be initialized to 0 to indicate the start of a page. This is shown by step 102, and is done at the start of each new scanned page.

Block 103 represents detection of a Group 4 code. As mentioned, there are three types of Group 4 codes – horizontal, vertical, and pass. Detecting Group 4 codes may be implemented by character string recognition, or other method from the wide variety within the ambit of one skilled in the art for code detection. Once a Group 4 code is detected, its type must be determined. That is, which of the three Group 4 types is it and, if it is a pass code, it is a white pass code or a black pass code. This process is done through a series of tests, the results of which determine how the x coordinate is determined. The x coördinate is referred to in terms of an old value of x, that is one that is associated with the code located by step 103 in the immediately-preceding pass through coordinate determination routine 100, and a new value of x, that is one that is associated with the code located by step 103 in the current pass through coordinate determination routine 100.

A code having been located, a test is performed to determine whether that code is a horizontal code. If so, the new value of x will be the old value of x added to the displacement value associated with the horizontal code. That is, the horizontal mode of Group 4 includes a code indicating the mode and a displacement indicating the number of pixels between the reference pixel color transition and the current pixel color transition. In the case of a horizontal code, the displacement is the number of pixels between a pixel color transition on the particular line and the next pixel color transition on that same line. This is indicated at 106.

It is important to note that this new value of x will not become an abscissa value used to determine alignment. Rather, this value is the running value of the displacement from the first pixel position on a scan line. Only the x values relating to white pass codes are used for alignment determination.

Proceeding for the moment with the assumption that a horizontal code is detected, a binary pixel color state bit is then incremented at 122, the method of and purpose for doing so are described in detail below. Once the new value of x has been calculated, x is checked at 108 to determine if the line end has been reached. This may be conveniently done by comparing x to the known length of a scan line, in pixels, and should x reach this value the line end has been detected. If the line end has not been reached, code detection continues for that line at 103. If the line end has been reached, x is set to 0 at 110 to correspond to the beginning of the next line and y, which keeps a running count of the line number, is incremented by one and checked at 111 to determine if the page end has been reached. This may be done, as with x, by comparing y to the known number of lines per page, and if y reaches this number a page end has been detected. If a page end has been reached, power is determined for various alignments swept through a number of alignment angles, at 126, as discussed in detail below. If page end has not been reached, code detection then resumes at 103.

If the code is determined not to be horizontal, it must be tested to determine if it is a vertical code, which is done at 112. It is necessary to store each transition position from the previous line for use as values on the reference line in the vertical mode. This is done in an array of the type $\{x1_r, x2_r, \ldots xn_r\}$, where r indicates a value on the reference line (the preceding scan line) and n is the total number of codes on the reference line. It is only necessary to preserve x values for the previous line, therefore, the new x values may overwrite the old once determined. This is shown at 114.

Assuming now that a vertical code has been detected, the value for x is then calculated at 116 as follows. The vertical mode is coded from a displacement between a pixel transition at pixel b_i on the reference line and a pixel transition at pixel a on the coding line. Since for the calculation of this displacement we are concerned only with x values, this displacement can be calculated simply as $|x_{b_i} - x_{a_i}|$, where x_{b_i} and x_{a_i} are the x values for pixels b_i and a_i, respectively. The new x value is the old x value added to this displacement. Again, the binary pixel color state bit is incremented at 122, and the new x value is tested at 108 to determine if the end

Note the three dashed arcs in Fig. 10. The innermost arc shows the power level to be expected if the fiducial points were unaligned. The next arc is drawn at the average power level over the 80 degree sweep. The outer-most arc is drawn at the peak power level. These arcs demonstrate the statistical significance of the peak power level as determined by the power calculation algorithm.

If a Baird-like alignment process is applied to the fiducial points which correspond to the positions of white passes in the Group 4 compressed representation of the image, an efficient and accurate means of determining the skew angle in the underlying image results.

Testing

Testing of the present system on CCITT test data demonstrates that the present invention provides a significant improvement in skew angle detection performance in a majority of cases. A comparison of skew angle detection performance between the method of Baird and the present method is shown in table 1 below. In each instance both skew angle detection algorithms were applied to the raw bitmaps to determine the amount of skew detected by the algorithm in the supposedly unskewed image. Then the bitmaps were digitally rotated through -3.0 degrees and +5.0 degrees, and the skew angle determination was performed again.

	No Rotation (degrees rotation)		-3.0 Degree Rotation (degrees rotation)		+ 5.0 Degree Rotation (degrees rotation)	
File	Baird	Present Invention	Baird	Present Invention	Baird	Present Invention
cc1	0.20	0.24	-2.76	-2.80	5.20	5.24
cc2	-1.00	0.10	-4.00	-2.12	6.00	5.02
cc3	-0.32	-0.32	-3.16	-3.24	4.88	4.76
cc4	-0.04	-0.08	-3.00	-3.00	4.92	5.00
cc5	0.52	0.48	-2.48	-2.52	5.52	5.48
cc6	0.08	0.24	-2.88	-2.88	5.08	5.12
cc7	0.16	0.16	-2.84	-2.88	5.08	5.12
cc8	0.08	-0.20	-3.20	-2.96	4.92	5.12

Table 1

One notable result of the above tests is that the step of classifying marks as being either text or non-text has been obviated. In fact, the present invention yields highly relevant and accurate results for images such as cc2, which are predominantly non-textual. This result is especially important at high skew angles, arising from the fiducial points being explicitly tied to the structure whose skew angle is being determined.

Presence of halftones with non-orthogonal screen angles may not degrade performance of the algorithm of the present invention since the halftone areas should be encoded using the uncompressed mode of the Group 4 recommendation and not involve the use of pass codes. Nevertheless the effect of the presence of halftone material on this technique has not yet been tested.

Characterizations of the shape of the power distribution can lead to useful generalizations about the predominance of a single skew angle in the page image. With reference to Fig. 11, the presence of multiple well-

of those pass codes as the locations of the selected points of pixel color transition.

8. The method of claim 7, including the further step of using those pass codes that indicate a pass of white pixels.

9. A method for determining skew angle or angles of a scanned image, wherein skew is determined by forming a count of selected topographic features of the image, forming selected weighted sums of the selected topographic features such that the topographic features form a plurality of alignments, and determining as the skew angle or angles that alignment or alignments which maximize one or more of the weighted sums, including the steps of:
   selecting a plurality of points located on the image, and
   using the selected points as the topographic features.

10. The method of claim 9, wherein the plurality of selected points on the image are line ends of the image.

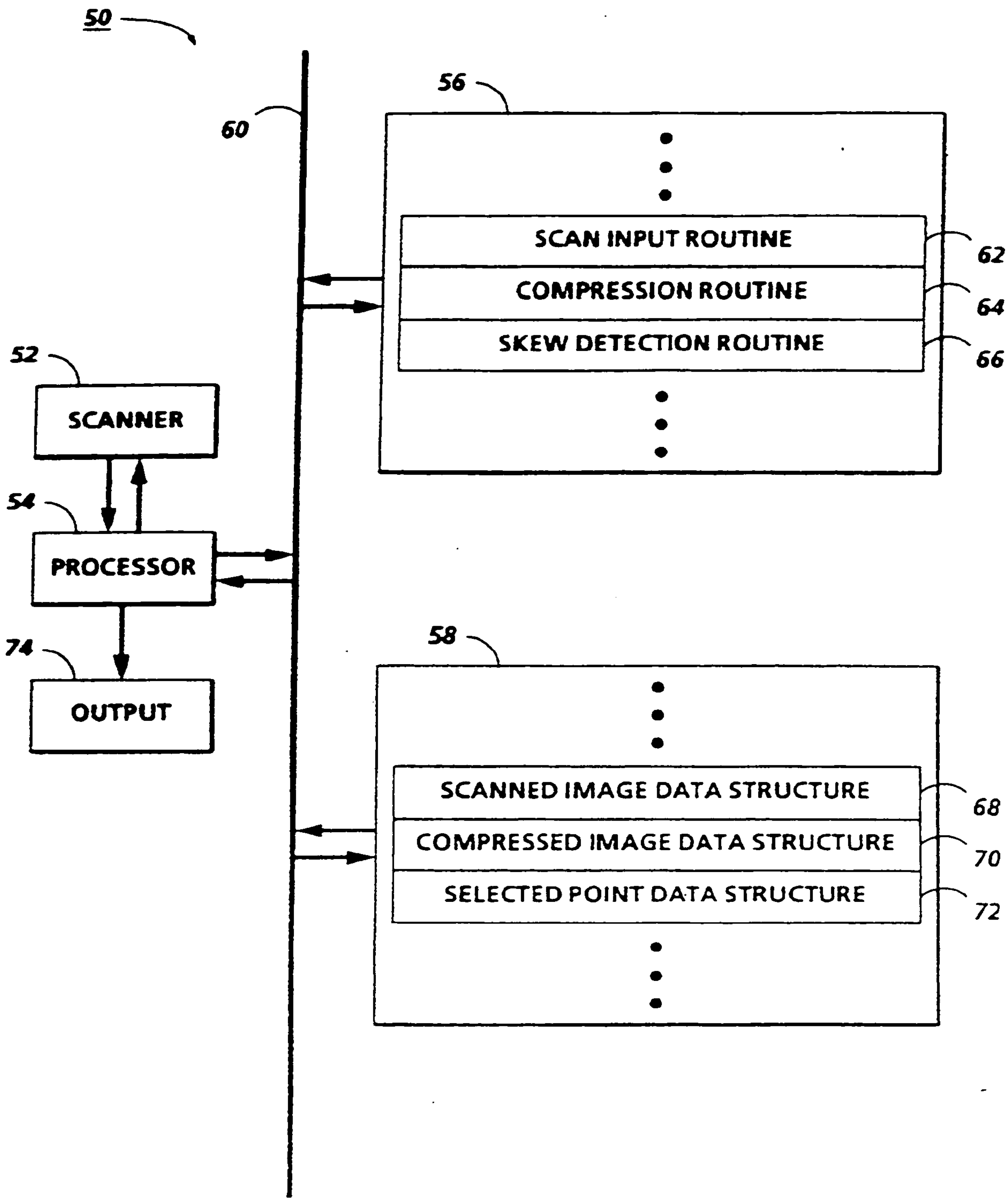

50
60
56
SCAN INPUT ROUTINE
62
COMPRESSION ROUTINE
64
SKEW DETECTION ROUTINE
66
52
SCANNER
54
PROCESSOR
74
OUTPUT
58
SCANNED IMAGE DATA STRUCTURE
68
COMPRESSED IMAGE DATA STRUCTURE
70
SELECTED POINT DATA STRUCTURE
72

FIG. 2

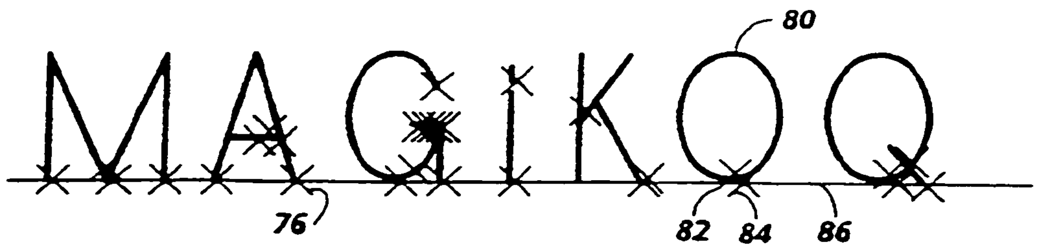


FIG. 4A

FIG. 4B

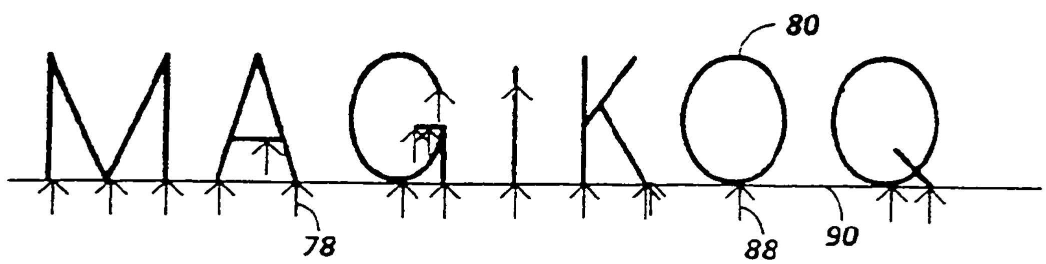


FIG. 5A

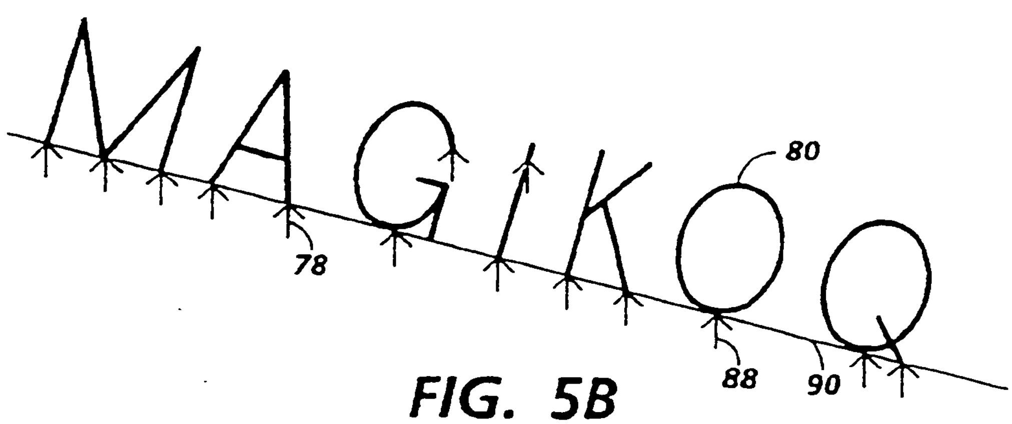


FIG. 5B

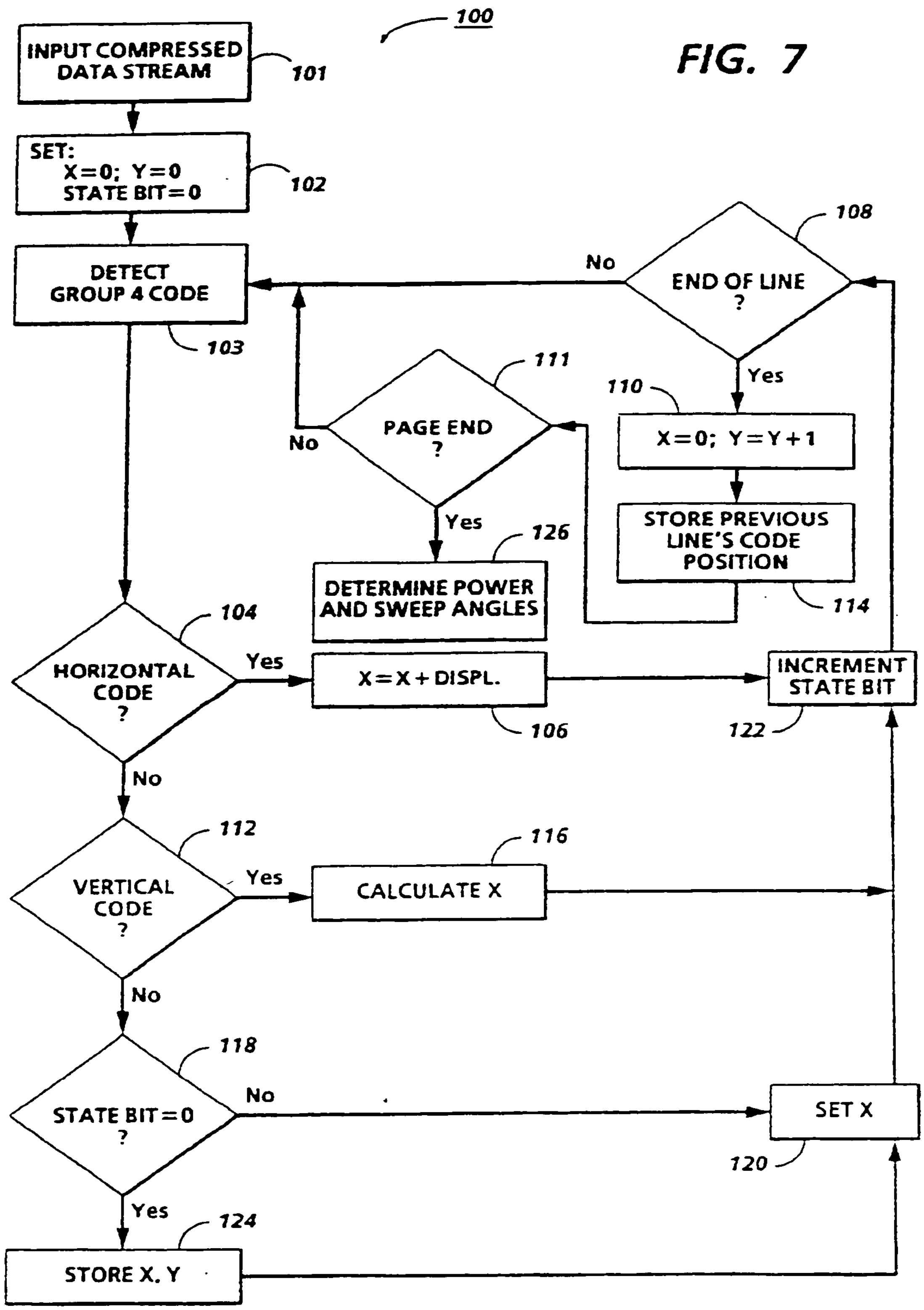

100
INPUT COMPRESSED DATA STREAM
101
SET: X=0; Y=0 STATE BIT=0
102
DETECT GROUP 4 CODE
103
END OF LINE ?
108
No
Yes
110
X=0; Y=Y+1
STORE PREVIOUS LINE'S CODE POSITION
114
PAGE END ?
111
No
Yes
126
DETERMINE POWER AND SWEEP ANGLES
HORIZONTAL CODE ?
104
Yes
X=X+DISPL.
106
INCREMENT STATE BIT
122
No
VERTICAL CODE ?
112
Yes
CALCULATE X
116
No
STATE BIT=0 ?
118
No
SET X
120
Yes
STORE X, Y
124

FIG. 7

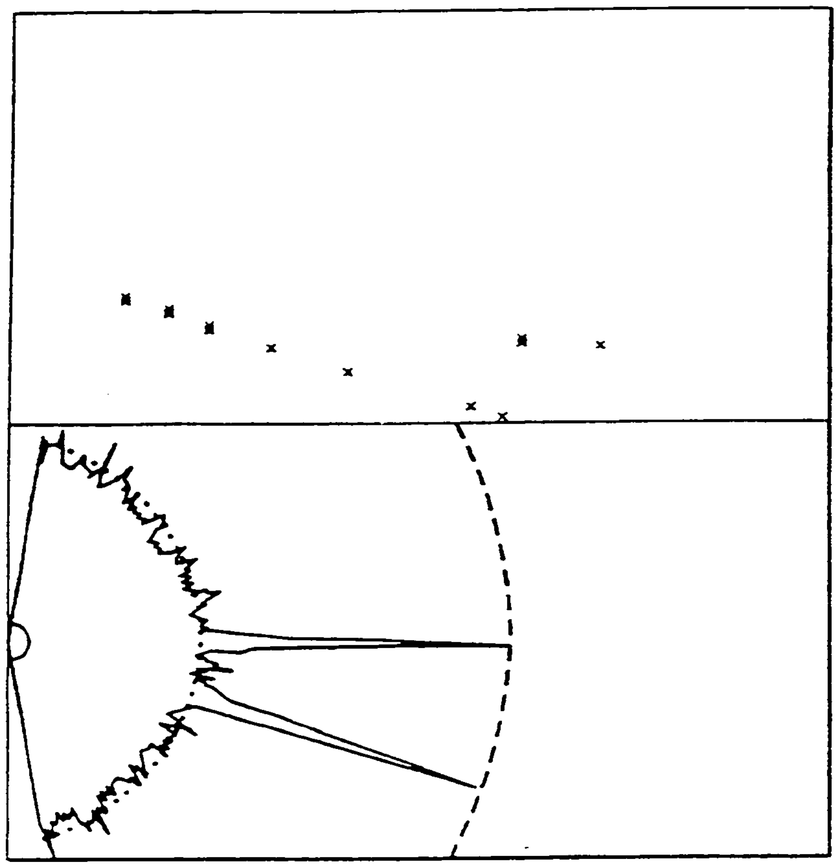


FIG. 11

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number : **0 434 415 A3**

(12) EUROPEAN PATENT APPLICATION

(21) Application number : 90313970.7

(22) Date of filing : 20.12.90

(51) Int. Cl.[5] : **G06K 9/32**

(30) Priority : 21.12.89 US 454339

(43) Date of publication of application :
26.06.91 Bulletin 91/26

(84) Designated Contracting States :
DE FR GB

(88) Date of deferred publication of search report :
21.11.91 Bulletin 91/47

(71) Applicant : XEROX CORPORATION
Xerox Square
Rochester New York 14644 (US)

(72) Inventor : Spitz, A. Lawrence
2311 Bryant Street
Palo Alto, California, 94301 (US)

(74) Representative : Weatherald, Keith Baynes et al
Rank Xerox Patent Department Albion House,
55 New Oxford Street
London WC1A 1BS (GB)

(54) Method of measuring skew angles.

(57) Skew angle of an image is determined based on determination of location of fiducial points on the image. Fiducial points may be located through a comparison of the scanning of a first line with scanning of a subsequent line. These fiducial points may be defined in terms of pixel color transitions located on a first scan line without a corresponding transition on the succeeding scan line. Skew angle may be determined from image data in uncompressed form or in compressed form. Where skew angle is determined from image data in compressed form, the two-dimensional CCITT facsimile recommendations may be used. In such cases, the locations of the fiducial points may be taken as the locations of the pass codes of the compressed image data. Specifically, pass codes indicating a pass of white pixels are used.

EP 0 434 415 A3

Jouve, 18, rue Saint-Denis, 75001 PARIS

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 90 31 3970

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
Y	WO-A-8 905 495 (EASTMAN KODAK COMPANY) * claim 1 *	1,3,4	G06K9/32
Y,D	PROCEEDINGS OF THE IEEE vol. 68, no. 7, July 1980, NEW YORK pages 854 - 867; ROY HUNTER ET AL.: 'INTERNATIONAL DIGITAL FACSIMILE CODING STANDARDS ' * page 864, left column, line 17 - page 865, left column, line 18; figure 6 *	1,3,4,5	
A		7,8,10	
Y,D	PROCEEDINGS OF SPSE SYMPOSIUM ON HYBRID IMAGING SYSTEMS 1987, pages 21 - 24; HENRY S. BAIRD: 'the skew angle of printed documents ' * page 22 *	5	
A	EP-A-308 673 (KABUSHIKI KAISHA TOSHIBA) * column 7, line 40 - column 10; figures 2,10,12 *	1,5,9	

TECHNICAL FIELDS SEARCHED (Int. Cl.5)
G06K
G06K
G06F
H04N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	05 SEPTEMBER 1991	CHATEAU J.P.

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)